

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **07244494 A**

(43) Date of publication of application: **19.09.95**

(51) Int. Cl

G10L 3/00

(21) Application number: **06034580**

(71) Applicant: **RICOH CO LTD**

(22) Date of filing: **04.03.94**

(72) Inventor: **MIYAMOTO KEIICHI**

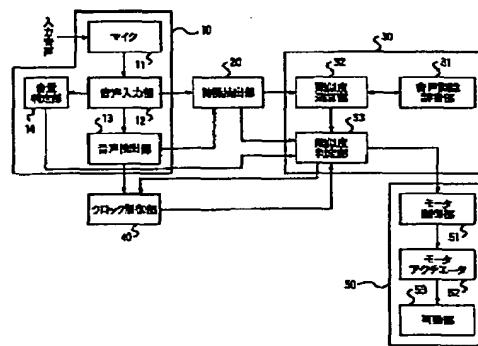
(54) SOUND RECOGNIZING DEVICE

(57) Abstract:

PURPOSE: To provide a sound recognizing device in which sound recognition can be performed even in an operation state of low power consumption.

CONSTITUTION: This device is a sound recognizing device which recognizes input sound, and provided with a similarity comparing and discriminating means 30 which performs comparing and discriminating operation of similarity between feature pattern data including feature quantity in a sound section in an input sound signal and the reference pattern data for plural sound previously stored based on a clock signal, and a control means 40 which lowers a frequency of the clock signal lower than the reference value when an input/output side is connected to the similarity comparing and discriminating means and a result of similarity comparing and discriminating is the prescribed sound to be expected, or when sound input of fixed time or more is not inputted.

COPYRIGHT: (C)1995,JPO



(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-244494

(43)公開日 平成7年(1995)9月19日

(51)Int.Cl.
G 10 L 3/00

識別記号 庁内整理番号
571 J

F I

技術表示箇所

審査請求 未請求 請求項の数4 OL (全5頁)

(21)出願番号 特願平6-34580

(22)出願日 平成6年(1994)3月4日

(71)出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72)発明者 宮本 恵一

東京都大田区中馬込1丁目3番6号 株式
会社リコー内

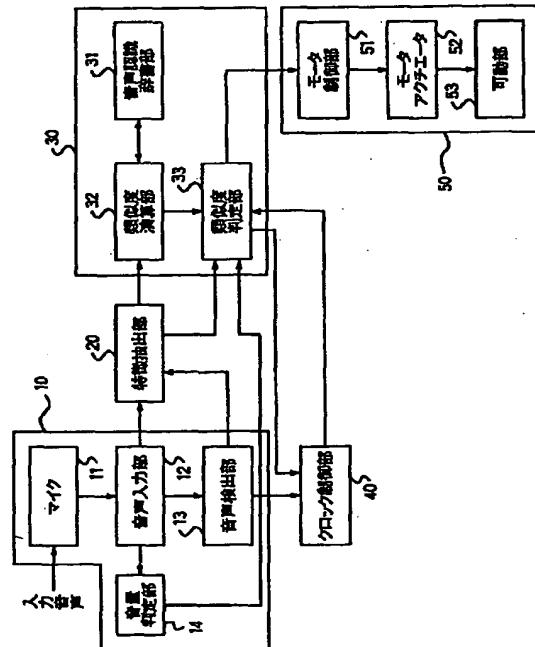
(74)代理人 弁理士 青山 葵 (外2名)

(54)【発明の名称】 音声認識装置

(57)【要約】

【目的】 低消費電力動作状態においても音声認識が可能な音声認識装置を提供する。

【構成】 入力音声信号における音声区間における特徴量を含む特徴パターンデータと複数の音声に対する予め格納済の基準パターンデータとの類似度の比較判定動作をクロック信号に基づき行う類似度比較判定手段30を備え、入力音声の認識を行う音声認識装置であって、入出力側が上記類似度比較判定手段に接続され上記類似度比較判定結果が所定の音声候補である場合若しくは一定時間以上音声入力がない場合には上記クロック信号の周波数を基準値よりも低下させる制御手段40を備えた。



1

【特許請求の範囲】

【請求項1】 入力音声が供給され入力音声信号における音声区間における特徴量を含む特徴パターンデータを送出する前処理手段と、上記前処理手段に接続され上記特徴パターンデータと複数の音声に対する予め格納済の基準パターンデータとの類似度の比較判定動作をクロック信号に基づき行う類似度比較判定手段とを備え、入力音声の認識を行う音声認識装置であって、
入力側が上記前処理手段及び上記類似度比較判定手段に接続されかつ出力側が上記類似度比較判定手段に接続され上記類似度比較判定結果が所定の音声候補でない場合若しくは一定時間以上音声入力がない場合には上記クロック信号の周波数を基準値よりも低下させる制御手段を備えたことを特徴とする音声認識装置。

【請求項2】 上記類似度比較判定手段は、周波数の低下したクロック信号が供給された場合には上記特徴パターンデータとの類似度比較判定の対象とする上記基準パターンデータを所定の基準パターンデータに限定する、請求項1記載の音声認識装置。

【請求項3】 類似度比較判定の対象とする上記基準パターンデータを所定の基準パターンデータに限定している状態において類似度比較判定結果が上記限られた基準パターン中で所定の確度を有するとき、上記類似度比較判定手段は上記制御手段に対して上記クロックの周波数を上記基準値へ増加させる信号を送出する、請求項2記載の音声認識装置。

【請求項4】 上記制御手段に対して上記クロックの周波数を上記基準値へ増加させる場合、上記類似度比較判定手段は以後の類似度比較判定対象を上記基準パターンデータのすべてとする、請求項3記載の音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声認識装置に関し、特に音声認識結果に基づき何らかの動作をするものであって低消費電力化を効果的に行う音声認識装置に関する。

【0002】

【従来の技術】 音声認識を利用した機器の中でも、携帯性を重視したものや、電池駆動を前提としたがん具等は、その音声認識を行う構成部分の消費電力が問題とされている。一般に音声認識を行う部分は、音声の特徴パターンを抽出する特徴抽出部分と、その特徴パターンがどの基準(標準)パターンと類似しているかを見出す類似度演算/判定部と、から構成されている。上記特徴抽出部分は、信号をアナログ処理する場合にはバンドバスフィルターバンク等、またディジタル処理を行う場合にはA/DコンバーターとDSP等で構成される。上記類似度演算/判定部は、CPU(中央演算処理装置)が類似度演算のアルゴリズムにのっとったソフトウェアで実行する場合や、専用のハードウェアロジックで行う場合が

2

多い。このような音声認識に係わる部分は一般的に消費電力が多く、常時動作させていると、電源である電池(一次、二次)の消耗が激しくなり、実用的でない。よってこのような欠点を克服するために種々の提案がなされている。

【0003】 例えば、特開昭62-245296号公報、特開平3-202899号公報(以下、これら公報の発明を従来技術1とする)には、一定期間音声の入力がないと音声認識部への電源供給を停止するものが開示されている。又、特開昭58-55991号公報(以下、この公報の発明を従来技術2とする)には従来技術1に加えて、常時動作している音声検出部が音声を検知したときに認識部への電源供給を開始することにより、音声認識機能を再開させようとするものが開示されている。

【0004】

【発明が解決しようとする課題】 従来技術1では電源が切れた後に音声認識動作をさせたいときは別途電源スイッチをオンする等の操作が必要であり煩わしさがある。

又、従来技術2では従来技術1の欠点を克服しているようであるが、実際には低消費電力の音声検出部としては上記公報に開示されるような非常に簡単な積分器程度の構成しか期待できないため、雑音による誤動作や、音声始端の欠落などがあり、その検出結果から切り出した音声区間を認識対象としたときの認識精度は高くない。また、区間検出後に電源が入る第1回目の発声では音声始端の欠落はより多くなってしまう。この欠落を避けるには、BBD等のアナログ遅延素子を採用すればよいが、もはや低消費電力を目指した構成とはならない。

【0005】 一方、玩具等のアプリケーションでは、音声認識部分の電源が切れている場合でも何らかの方法にて音声で動作させたほうが望ましい場合があるが、従来技術1の構成ではこれに対応できない。又、従来技術2の構成では通常動作時の高い認識精度が期待できず、又、低消費電力化が困難である。本発明は上述したような問題点を解決するためになされたもので、低消費電力動作状態においても音声認識が可能な音声認識装置を提供することを目的とする。

【0006】

【課題を解決するための手段】 本発明は、入力音声が供給され入力音声信号における音声区間における特徴量を含む特徴パターンデータを送出する前処理手段と、上記前処理手段に接続され上記特徴パターンデータと複数の音声に対する予め格納済の基準パターンデータとの類似度の比較判定動作をクロック信号に基づき行う類似度比較判定手段とを備え、入力音声の認識を行う音声認識装置であって、入力側が上記前処理手段及び上記類似度比較判定手段に接続されかつ出力側が上記類似度比較判定手段に接続され上記類似度比較判定結果が所定の音声候補でない場合若しくは一定時間以上音声入力がない場合

50

には上記クロック信号の周波数を基準値よりも低下させる制御手段を備えたことを特徴とする。

【0007】

【作用】このように構成することで制御手段は、例えば一定時間以上音声の入力がない場合にはクロック信号の周波数を基準値より低下させることで、類似度比較判定手段の動作速度を下げ消費電力の低減を図るように作用する。又、類似度比較判定手段は、低消費電力時にあっても音声認識を行うように作用する。

【0008】

【実施例】本発明の音声認識装置の一実施例を図を参照しながら以下に説明する。図1に示すように本実施例の音声認識装置は以下の構成をなす。本実施例の音声認識装置は、音声に従い何らかの動作を行う例えがん具に使用されるもので、大きく分類すると入力音声の前処理部10と、入力音声信号から特徴パターンを抽出する特徴抽出部20と、特徴パターンと基準パターンとの類否を演算判定する類似度演算判定部30と、類似度演算判定部30へ供給するクロック信号の周波数を制御するクロック制御部40と、動作部50とに区分される。

【0009】前処理部10は、マイク11と、マイク11から音声信号が供給される音声入力部12と、音声入力部12の出力側に接続され音声区間を検出し該音声区間の開始及び終了に対応する時間情報を送出する音声検出部13と、入力側が音声入力部12に接続され音声の音量を判定する音量判定部14とを備える。又、音声検出部13の出力側は特徴抽出部20及び後述のクロック制御部40に接続され、音声入力部12の出力側は特徴抽出部20に接続され、音量判定部14の出力側は後述の類似度判定部33に接続される。

【0010】特徴抽出部20は、音声検出部13から供給される上記時間情報に基づき、音声入力部12から供給される音声信号から音声区間を選択し、該音声区間について例えば10m秒毎に音声信号の特徴量を抽出する。特徴抽出部20の出力側は、類似度演算部32及び類似度判定部33に接続される。

【0011】類似度演算判定部30は、音声認識辞書部31と、類似度演算部32と、類似度判定部33とを備えている。音声認識辞書部31は、複数の音声のそれについての特徴パターンを格納している。尚、特徴パターンとは、上述したように例えば特徴抽出部20のような装置にて抽出された音声の特徴量を集合したものである。類似度演算部32は、入力側に特徴抽出部20の出力が接続されかつ音声認識辞書部31と入出力が接続され、特徴抽出部32から供給される特徴量を音声区間分集合させてなる特徴パターンと音声認識辞書部31から供給される特徴パターンである基準パターンとを比較し類似度を演算する。類似度判定部33は、類似度演算部32から供給される類似度情報について所定のしきい値を基準にして入力音声が音声認識辞書部31に格納す

るいづれの単語と一致又は近似するかを判別する。類似度判定部33の出力側は、クロック制御部40及び後述のモータ制御部51に接続される。

【0012】クロック制御部40の出力側は類似度演算判定部30の各構成部分31ないし33に接続され、クロック制御部40は、類似度演算判定部30の各構成部分31ないし33に対して選択的にクロック周波数を変化させ、供給又は非供給するものであり、音声認識結果が所定の候補である場合もしくは、一定時間以上音声入

10 力がなかった場合にはクロック周波数を基準値よりも低下させる。また、この状態において認識結果が前記限られた一部の認識辞書中であらかじめ定められた確度以上であった場合、クロック周波数を上記基準値まで増加させるように前記クロック制御部を動作するように成し、同時にそれ以降の認識対象を前記音声認識辞書全体にするように成したことと特徴としている。

【0013】動作部50は、可動部53と、該可動部53を動作させるモータ/アクチュエータ52と、モータ/アクチュエータ52の動作を制御するモータ制御部51と20 を備え、モータ制御部51の入力側には類似度判定部33の出力が接続される。

【0014】このように構成される本実施例における音声認識装置の動作を以下に説明する。尚、本実施例の音声認識装置は図2に示すようなロボットおもちゃに備わるとしている。音声認識辞書31には、図3に示すように、予め数語の音声認識特徴量(以下パターン)が登録されているとする。本例においてその特徴パターンは、「めざめよ」、「スター」、「進め」、「やっつけろ」とあるとする。これらの特徴パターンには順に1,2,3,4と単語番号が付される。このとき単語番号「1」には目覚めの動作、単語番号「2」にはそのロボットの名前、単語番号「3」には前進に対応する言葉、単語番号「4」には攻撃動作に対応する言葉というように、予め大体の意味内容との対応が決められているとする。

【0015】まず、音声の一般的な認識動作を説明する。マイク11から入力された操作者の音声は、音声入力部12を介して音声検出部13へ供給されるとともに特徴抽出部20へ供給される。音声検出部13では、入力音声情報から音声と判断される区間を選択的に選び出し(区間検出:セグメンテーション)、その区間の開始及び終了に関する時間情報を送出する。特徴抽出部20

40 は、音声入力部12から供給される音声信号と、音声検出部13から供給される上記区間の開始、終了に関する時間情報とに基づき、音声が存在する区間から特徴量を抽出し、該特徴量を類似度演算部32へ送出する。

【0016】該特徴量は類似度演算部32において音声区間分集合され特徴パターンとなり、該特徴パターンは音声認識辞書31内のそれぞれの基準パターンと照合され、各特徴パターン毎に類似度が計算される。類似度判定部33では、音量判定部14から供給される音量情報

と類似度演算部32から供給される、各特徴パターンに対する入力音声の類似度情報に基づき、入力音声が音声認識辞書31に格納された単語のいずれに最も近いかを判定し、該判定結果に基づきモータの動かしかたを決定する。

【0017】モータ制御部51では、類似度判定部33から供給されるモータの動かしかたの情報を選択し、モータ/アクチュエータ52を駆動させる。これによって可動部53は動作する。尚、本実施例では、可動部53は図2に示すロボットの腕、首、足である。類似度判定部33における認識結果が「バスター」ならば、ロボットの全身をがたがた震わせる、「進め」ならロボットの足を動かして前進し、「やっつけろ」ならロボットの腕を振る、等と動作を対応付けておけば楽しい玩具となろう。

【0018】次に、本実施例の音声認識装置におけるクロック制御について説明する。主電源(不図示)が投入された直後には当該音声認識装置全体の動作を制御するシステムコントローラ(不図示)から送出される制御信号にて、又、予め定められた期間以上発声者が発声しない場合には音声検出部13から上記時間情報の供給がないことを検知することで、クロック制御部40は予め定めた、基準値よりも低い動作周波数(以下、低クロックと記す)を類似度演算判定部30を構成する各構成部分31ないし33へ送出する。又、同時に、類似度演算部32は低クロックの供給を検知することで、認識対象単語について単語番号1、2のみに限定する。このように類似度演算部32は、低クロックの供給を検出したときは音声認識辞書31に格納される第1番目の単語から所定数のみを認識対象とすることもできるが、音声認識辞書31に格納される各単語の基準パターンを示すデータに低クロック状態には認識対象となる旨の所定データを予め付加しておき、上記所定データの有無を判断することで認識対象の単語を選択するようにしてよい。

【0019】このように認識対象を限定することで低クロック時でも認識処理に要する時間が長大化することを防ぐことができる。即ち、低クロック供給状態では、類似度演算部32に供給された音声情報は、単語番号

「1」の「めざめよ」と、単語番号「2」の「バスター」との2単語のみと類似度が演算される。この演算の結果、類似度判定部33は、認識の確からしさを表す、確度が予め定められたある閾値以上である場合、即ち、入力音声を上記単語番号「1」、「2」のいずれかであると認識できた場合にはクロック制御部40に対して制御信号を送出する。クロック制御部40は、上記制御信号の供給に基づき、上記基準値である予め定められた高い動作周波数(以下、高クロックと記す)を類似度演算判定部30の各構成部分31ないし33へ送出し始める。そして類似度演算部32は、高クロックの供給を判断材料として、以後の認識対象単語を音声認識辞書31に格納するすべての単語にするように制御する。尚、上記確度が上

記閾値に満たない場合、即ち入力音声を認識できなかった場合には、類似度判定部33は、基本的に、クロック制御部40に対して制御信号を送出することはなくそのまま低クロックにて動作するが、認識できなかったとはいえ何等かの音声が入力されたことに変わりはないことから、上記制御信号を出し高クロックにて動作するようにしてよい。又、類似度判定部33が送出する認識結果に応じて、入力音声を認識したことを音声発生者に知らせるために入力音声に対して例えば応答するような所定の動作を動作部50が完了した後に類似度演算部32は以後の認識対象単語を音声認識辞書31に格納するすべての単語とするようにしてよい。

【0020】したがって高クロック供給開始以降は、N(Nは自然数)単語例えは40単語を認識対象としても十分に速い反応時間を達成することができる。この場合認識処理のみに着目すれば、対象単語は低クロック供給時で2単語、高クロック供給時で40単語であるから、供給するクロックの周波数比は2:40とすることができるところがわかる。LSI(大規模集積回路)の主流であるCMOSプロセスでは、クロック周波数と消費電流とがほぼ比例するため、低クロック供給時における周波数を下げることで消費電力をより低減することができその効果は大きい。又、認識動作を繰り返した後、一定期間音声入力がなければ、類似度判定部33からクロック制御部40へ送出される制御信号により再びクロック制御部40は低クロックを送出し、消費電力を低減する。

【0021】尚、上述した実施例では、低クロック送出の基準として音声入力のない時間を材料としたが、キーワードを決めておきこのキーワードが認識結果となったときに、低クロック送出状態に移行するようにしてよい。このようにキーワードを使用した場合、例えば上記キーワードを「眠れ」とすると、「眠れ」の音声を認識したときには低クロック状態となり低消費電力状態とするようなことが可能となり、例えばがん具に本実施例の音声認識装置を使用した場合にはよりおもしろい態様を得ることができる。

【0022】又、本実施例では音声認識のアルゴリズムや特徴量また可動部および電源制御部の制御方法等について何ら限定するものではない。

【0023】【発明の効果】以上詳述したように本発明によれば、例えば一定時間以上音声の入力がない場合にはクロック信号の周波数を基準値よりも低下させるようにしたことにより、類似度比較判定手段の動作速度を下げ消費電力の低減を図ることができる。又、このような低消費電力時においても類似度比較判定手段にはクロック信号が供給されているので、音声認識を行うことができる。

【図面の簡単な説明】

【図1】 本発明の音声認識装置の一実施例における構成を示すブロック図である。

【図2】 図1に示す音声認識装置が使用される一例であるがん具を示す図である。

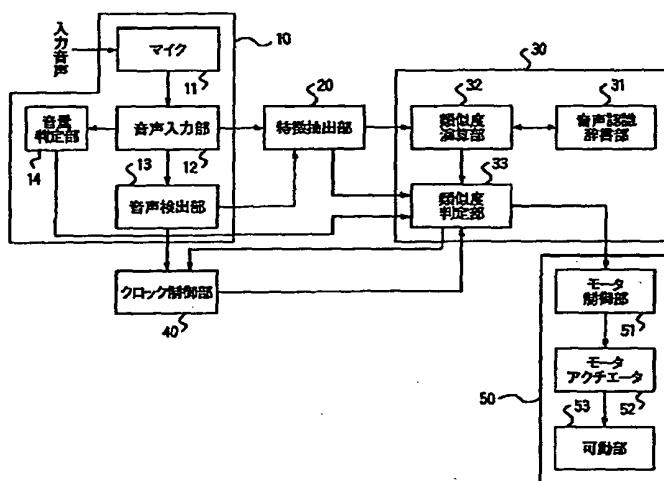
【図3】 図1に示す音声認識辞書に格納される単語等を示す図である。

【符号の説明】

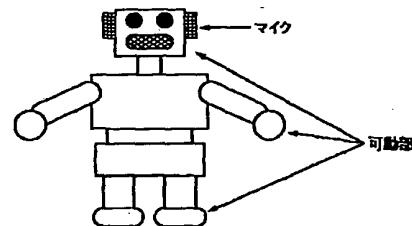
10…前処理部、12…音声入力部、13…音声検出 *

*部、14…音量判定部、20…特徴抽出部、30…類似度演算判定部、31…音声認識辞書部、32…類似度演算部、33…類似度判定部、40…クロック制御部、50…作動部、51…モータ制御部、52…モータ/アクチュエータ、53…可動部。

【図1】



【図2】



【図3】

単語番号	単語音声	動作
1	めざめよ	ガオーと音を出し頭を上げる
2	バスター	全身をがたがた震わせる
3	進め	足を動かして前進する
4	やっつけろ	頭を振る
5	(以下略)	(以下略)
6		
・		
・		
・		
N-1		
N		

